Supplementary III:

In this study, a total of 1542 radiomic features were extracted in this study. Here, we will give more details about the features.

Gabor filter

Gabor filters detect edges through response to Gabor wavelet features. Each pair *(f, θ)* shows a specific Gabor filter at a specific frequency *(f = 0, 2, 4, 8, 16, 32)* and orientation *(θ = 0, π/8, π/4, 3π/8, π/2, 5π/8, 3π/4, 7π/8)*. The Gabor filter is basically a Gaussian (with variances sx and sy along x and y-axes respectively), modulated by a complex sinusoid (with center frequencies U and V along x and y-axes respectively) described by the following equation:

$$G(x, y, \theta, f) = \exp\left(-0.5\left[\left(\frac{x'}{sx'}\right)^2 + \left(\frac{y'}{sy'}\right)^2\right]\right) * \cos(2\pi f x')$$

Where x' and y' are defined as follow:

$$x' = x * \cos(\theta) + y * \sin(\theta)$$
$$y' = y * \cos(\theta) - x * \sin(\theta)$$

The histogram contains 48 bins corresponding to different angles and frequencies. The block normalization is then accomplished by the L2 norm normalization of the vector. Each Gabor filter is then convolved with the original image and values corresponding to filter response within the ROIs are concatenated.

Laws filter

The Laws features may possibly detect patterns of heterogeneous enhancement and abnormal structure. The Laws filters response to 5 × 5-pixel filter combination of specific textural patterns. Features include all combinations of five 1D filters: level, edge, spot, wave, and ripple.

L (Level) = [1 4 6 4 1]

E (Edge) = [-1 -2 0 2 1]

S (spot) = [-1 0 2 0 -1]

W (wave) = [-1 2 0 -2 1]

R (Ripple) = [1 -4 6 -4 1]

To obtain a feature vector, each filter is convolved with the image and the absolute value of filter response within all voxels contained within a region of interest are concatenated.

Laws Laplacian features

Laplacian pyramids allow capturing multi- scale edge representations via a set of band pass filters. To do that the original image is first convolved with a Gaussian kernel. The Laplacian is then computed as the difference between the original image and the low-pass-filtered image. The resulting image is then sub- sampled by a factor of 2, and the filter subsample operation is repeated recursively. This process is continued to obtain a set of band pass-filtered images. Laws Energy filters are then applied to the resulting images to obtain a set of 25 features.

Haralick feature

Haralick features can capture the quantify heterogeneity and entropy of local intensity texture as represented by the gray-level co-occurrence matrix within a 5 × 5-pixel window. If $p(i,j)$ is a Gray level co-occurrence Matrix (GLCM) constructed from original image, then 13 Haralick features are defined as below.

1. Entropy $Ent = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p(i,j)\log\big(p(i,j)\big)$
2. Energy $E = \sqrt{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p^2(i,j)}$
3. Inertia (Contrast) $inertia = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i-j)^2\, p(i,j)$
4. Inverse Difference Moment $idm = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{1}{1+(i-j)^2}p(i,j)$
5. Correlation $corr = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{ijp(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y}$
6. Info. Measure of Correlation 1 $info1 = \frac{HXY-HXY1}{\max(HX,HY)}$
7. Info. Measure of Correlation 2 $info2 = \big(1-\exp\big(-2(HXY2-HXY)\big)\big)^{\frac{1}{2}}$
8. Sum Average $sum_{av} = \sum_{i=2}^{2N} i\, p_{x+y}(i)$
9. Sum Variance $sum_{var} = \sum_{i=2}^{2N}(i-f_8)^2\, p_{x+y}(i)$
10. Sum Entropy $-\sum_{i=2}^{2N} p_{x+y}(i)\log(p_{x+y}(i))$
11. Difference Average $diff\_av = \sum_{i=2}^{2N}(i-f_8)^2\, p_{x-y}(i)$

12. Difference Variance $diff_{var} = E\left[{p_{x-y}}^2\right] - E[p_{x-y}]^2$

13. Difference Entropy $diff_{ent} = -\sum_{i=2}^{2N} p_{x-y}(i)\log(p_{x-y}(i))$

$p_{x+y}(k) = \sum_{i=0}^{N-1}\sum_{i=0}^{N-1} p(i,j)$ $\quad i + j = k, k = 0,1,2, \ldots$

$p_{x-y}(k) = \sum_{i=0}^{N-1}\sum_{i=0}^{N-1} p(i,j)$ $\quad |i - j| = k, k = 0,1,2, \ldots$

$$p_x(i) = \sum_{j=0}^{N-1} p(i,j)$$

$$p_y(j) = \sum_{i=0}^{N-1} p(i,j)$$

$$HXY = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p(i,j)\log\big(p(i,j)\big)$$

$$HXY1 = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p(i,j)\log\big(p_x(i)p_y(j)\big)$$

$$HXY2 = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p_x(i)p_y(j)\log\big(p_x(i)p_y(j)\big)$$

Gray-Level Co-Occurrence Matrix

A GLCM is defined as $P(i,j;\sigma,\theta)$, a matrix with size N x N describing the second-order joint probability function of an image, where the (i,j)th element represents the number of times the combination of intensity levels i and j occur in two pixels in the image, that are separated by a distance of δ pixels in direction θ, and N is the number of discrete gray level intensities.

Co-occurrence of Local Anisotropic Gradient Orientations (CoLIAGe) features

CoLIAGe involves extracting the dominant gradient orientation along the X and Y directions for every pixel via principal component analysis. A co-occurrence matrix is computed for every pixel within the neighborhood (5 x 5) to capture co-occurring arrangements of the dominant gradient orientations. CoLIAGe features detects local differences in voxel-level gradient orientations for distinguishing similar appearing phenotypes [2].

HOG feature

In the HOG feature descriptor, the distribution (histograms) of directions of gradients (oriented gradients) are used as features. The image is divided first into several patches. To calculate a HOG descriptor, the horizontal and vertical gradients are calculated; This is easily achieved by filtering the image with the horizontal and vertical kernels of [-1, 0, 1] (Sobel operator). The magnitude and direction of gradient are then calculated. The image is then divided into 8×8 cells and a histogram of gradients is calculated for each 8×8 cells. The histogram contains 9 bins corresponding to angles 0, 20, 40 … 160. The contributions of all the pixels in the 8×8 cells are added up to create the 9-bin histogram. A 16×16 block normalization is then accomplished by the L2 norm normalization of vector. To calculate the final feature, vector for the entire image patch are concatenated into one giant vector.

In this study, HOG features were highly stable on test/retest dataset. The average and variance ICC for the HOG features was 0.88 ± 0.085, suggesting a high degree of reproducibility. However, from 20 HOG features, just one was significantly difference between responders and non-responders (p-value = 0.036). The MLR classifier trained with this feature in training set could discriminate responders from non-responders with an AUC of 0.68 in the test set.

LBP feature

The LBP tests the correlation between pixel and its neighbors, encoding this relation into a binary word. This allows detection of patterns, while being immune to contrast changes. The LBP features are computed using N sampling points on a circle of radius R and using mapping table defined by MAPPING. For each pixel in a cell, the pixel to each of its 8 neighbors is compared. Where the center pixel's value is greater than the neighbor's value, the pixel will be filled with "0", otherwise "1". This gives an 8-digit binary number (which is usually converted to decimal). The histogram, over the cell, of the frequency of each "number" is computed and then this histogram is normalized. This gives a feature vector for the entire window.

Shape feature

The surface area, volume, compactness 1 and 2, surface to volume ratio, sphericity, max diameter, major axis, minor axis, elongation, flatness, width, height, depth, perimeter, area, eccentricity, compactness, radial distance and roughness were calculated for nodules based on the definitions that can be found in the Computational Environment for Radiotherapy Research (CERR) [1].

References

[1]. J.O. Deasy, A.I. Blanco, V.H. Clark, CERR: a computational environment for radiotherapy research Med Phys, 30 (2003), pp. 979-985.

[2]. P. Prasanna, P. Tiwari & A. Madabhushi. Co-occurrence of Local Anisotropic Gradient Orientations (CoLIAGe): A new radiomics descriptor. Sci Rep. 2016; 6: 37241.